

AUDIO BUFFER FOR SYSTEM TARGET DECODER

Patent number: JP2000057691

Publication date: 2000-02-25

Inventor: FUJINAMI YASUSHI

Applicant: SONY CORP

Classification:

- international: **H04N5/44; H04N5/60; H04N7/24; H04N7/52; H04N7/62; H04N9/804; H04N5/04; H04N5/85; H04N9/806; H04N9/82; H04N5/44; H04N5/60; H04N7/24; H04N7/52; H04N9/804; H04N5/04; H04N5/84; H04N9/82; (IPC1-7): G11B20/10**

- european: H04N5/44N; H04N5/60N; H04N7/24C2; H04N7/52A; H04N7/62; H04N9/804B

Application number: JP19980226032 19980810

Priority number(s): JP19980226032 19980810

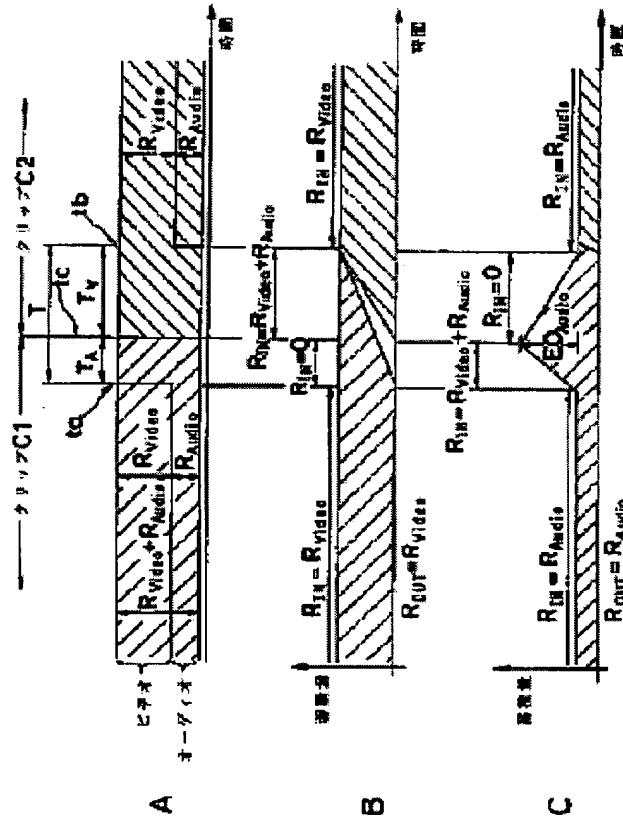
Also published as:

US6718119 (B1)

[Report a data error here](#)

Abstract of JP2000057691

PROBLEM TO BE SOLVED: To stipulate the capacity of an audio buffer to be increased in a system target decoder performing the seamless reproduction of a stream obtained by multiplexing an audio skew generated at the connection part of the stream so as not to be away from a video stream corresponding thereto. **SOLUTION:** The capacity of the audio buffer is increased by at least the following quantity: $T \times R_{VIDEO} + R_{AUDIO}$. In the formula, T is the sum of a period when an audio stream is inputted to the audio buffer, but a video stream is not inputted to a video buffer and a period when the audio stream is not inputted to the audio buffer, but the video stream is inputted to the video buffer, and R_{VIDEO} and R_{AUDIO} are respectively the read-out rates of the video stream and the read-out rate of the audio stream.



Data supplied from the esp@cenet database - Worldwide

【特許請求の範囲】

【請求項1】ストリームの接続部分で発生するオーディオスキューを、前記オーディオスキューに対応するビデオストリームと離れないように多重化したストリームのシームレス再生を行なうシステムターゲットデコーダーのオーディオバッファにおいて、前記オーディオバッファにオーディオストリームが入力されるが前記システムターゲットデコーダーのビデオバッファにビデオストリームが入力されない期間のデータ蓄積量の増加量に応じて、容量を増大させたことを特徴とするシステムターゲットデコーダーのオーディオバッファ。

【請求項2】請求項1に記載のシステムターゲットデコーダーのオーディオバッファにおいて、前記データ蓄積量の増加量を

$$T \times R_{\text{Video}} \times R_{\text{Audio}} \div (R_{\text{Video}} + R_{\text{Audio}})$$

(但し、Tは前記オーディオバッファにオーディオストリームが入力されるが前記ビデオバッファにビデオストリームが入力されない期間との前記オーディオバッファにオーディオストリームが入力されないが前記ビデオバッファにビデオストリームが入力される期間との和、 R_{Video} , R_{Audio} はそれぞれビデオストリーム、オーディオストリームの読み出しレート)として算出することを特徴とするシステムターゲットデコーダーのオーディオバッファ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、システムターゲットデコーダーの技術分野に属し、特に、ストリームの接続部分で発生するオーディオスキューをそれに対応するビデオストリームとは離れないように多重化したストリームのシームレス再生を行なうシステムターゲットデコーダーにおけるオーディオバッファの容量に関する。

【0002】

【従来の技術】MPEG2システム規格では、パケット多重化したストリームから個々のエレメンタリーストリーム(ビデオストリームやオーディオストリーム)を分離し、それらを復号して同期させて表示する過程を規定した仮想的なデコーダモデルとしてのSTD(システムターゲットデコーダ)を、PS(プログラムストリーム), TS(トランスポートストリーム)のそれぞれについて定義している。しかし、MPEG2システム規格には、複数のストリームを接続してシームレスに再生する方法についての規定はない。

【0003】そこで、DVD(ディジタルバーサタイルディスク)についての規格のうちのDVD-Video規格では、ユーザーがディレクターズカットやマルチアングルといった操作を行なえるようにすることを目的として、こうしたシームレス再生機能を導入した拡張デコーダモデルとしてのE-STD(エクステンデッドシ

テムターゲットデコーダ)を定義している。

【0004】シームレス再生用のストリームは、インターリーブユニットと呼ぶ単位に分割してインターリーブブロックと呼ぶ記録領域に配置される。そして再生時には、それらのインターリーブユニットのうちユーザーに指定されないものはジャンプされ、ユーザーに指定されたものだけが読み出されてE-STDに連続的に入力される。

【0005】このE-STDでは、ストリームの接続部分で発生するオーディオスキューを、それに対応する(同期再生すべき)ビデオストリームとは離して多重化する記録方式を前提としていた。即ち、例えば図8に示すように、或るストリームS1に続くストリームとして、マルチアングルのためにAngle1, Angle2, Angle3の3つのインターリーブユニットが用意される場合には、ストリームS1とAngle1, 2, 3との接続部分で発生するオーディオスキューを、ストリームS1中のビデオストリームとは離してAngle1, 2, 3のほうに多重化していた。

【0006】

【発明が解決しようとする課題】しかし、このような記録方式では、図8におけるAngle1, Angle2, Angle3中に、復号や表示のための基準時刻が互いに異なるデータが混在してしまうので、その論理構造が不明確になってしまふという問題を生じる。

【0007】また、ストリームの境界とは異なる箇所でストリームが切り離されるので、この切り離された1まとまりのデータ(図8におけるストリームS1やAngle1, 2, 3)のことを「クリップ」と呼ぶことになると、図9にも示すように、クリップの境界とストリームの境界とが一致しなくなることにより、境界の意味が不明確になるという問題を生じる。特に、個々のクリップをファイルとしてパーソナルコンピュータ等で管理する場合には、オーディオスキューとそれに対応するビデオストリームとが別々のファイル内に含まれてしまうことになるので、ファイル管理に支障が生じかねない。

【0008】これに対し、ストリームの接続部分で発生するオーディオスキューを、それに対応するビデオストリームとは離れないよう多重化する記録方式を採用すれば、こうした問題を回避することができる。

【0009】しかし、後に理由を説明するように、こうした記録方式のストリームのシームレス再生を行なうためのシステムターゲットデコーダとしては、MPEG2システム規格で定義されるP-STDのハードウェア資源のうちのオーディオバッファの容量を増加させたものか、あるいは、このハードウェア資源はそのままにして、单一のバッファのうちビデオバッファとして割り当てる領域とオーディオバッファとして割り当てる領域とを時間的に変化させるソフトウェア処理を行なうようにしたものが必要になる。

【0010】本発明の目的は、このうちのオーディオバッファの容量を増加させたほうのシステムターゲットデコーダにおける、増加すべきオーディオバッファの容量を規定することにある。

【0011】

【課題を解決するための手段】本発明は、ストリームの接続部分で発生するオーディオスキューをそれに対応するビデオストリームと離れないよう、多重化したストリームのシームレス再生を行なうシステムターゲットデコーダのオーディオバッファにおいて、このオーディオバッファにオーディオストリームが入力されるが、このシステムターゲットデコーダのビデオバッファにビデオストリームが入力されない期間のデータ蓄積量の増加量に応じて、容量を増大させたことを特徴としている。

【0012】本発明によれば、オーディオバッファの容量を上記の量だけ増大させたことにより、ストリームの接続部分で発生するオーディオスキューをそれに対応するビデオストリームと離れないよう、多重化したストリームのシームレス再生を行なうシステムターゲットデコーダシステムに、オーディオバッファをオーバーフローさせることなくシームレス再生を行わせることが可能になる。

【0013】

【発明の実施の形態】最初に、本発明が前提とする記録方式について説明する。この記録方式は、シームレス再生すべき2つのクリップのうちの先のクリップ（以下クリップC1と呼ぶ）中のオーディオストリームの最後のパケットに、クリップC1中のビデオストリームのうちの最後のピクチャの表示の終了時と一致した表示時刻のサンプルを含め、且つ、この2つのクリップのうちの後のクリップ（以下クリップC2と呼ぶ）中のオーディオストリームの最初のパケットに、クリップC2中のビデオストリームのうちの最初のピクチャのパケットの表示の開始時と一致した表示時刻のサンプルを含めることにより、ストリームの接続部分で発生するオーディオスキューをそれに対応するビデオストリームと離れないよう、多重化するものである。

【0014】この記録方式（以下では「クリーンブレーク」と呼ぶことにする。）を、従来の記録方式を例示した図8と関連付けて示すと、図1の通りである。このクリーンブレークによれば、1つのクリップ中には、復号や表示のための基準時刻が共通なデータのみが存在するようになるので、その論理構造が明確になる。

【0015】また、ストリームが境界で切り離されたものがクリップとなるので、図2にも示すようにクリップの境界とストリームの境界とが一致することにより、境界の意味が明確になり、クリップをファイル管理する場合にもオーディオスキューとそれに対応するビデオストリームとが同一のファイル内に含まれるようになる。

【0016】次に、本発明を適用したシステムターゲッ

トデコーダの構成の一例について、図3を用いて説明する。このシステムターゲットデコーダは、ディジタルビデオテープレコーダ用のシステムターゲットデコーダであり、以下「DVR-STD」と呼ぶこととする。

【0017】図示しないビデオテープから読み出されてDVR-STDに入力されたパケット多重化ストリーム（前述のクリーンブレークによるストリーム）は、デマルチプレクサ1により、エレメンタリーストリームであるビデオストリームとオーディオストリームとサブピクチャストリーム（グラフィックデータ等）とに分離されて、ビデオバッファ2、オーディオバッファ3、グラフィックバッファ4にそれぞれ蓄積される。

【0018】DTS（デコードタイムスタンプ）の示す復号時刻になると、ビデオストリームのうちのIピクチャ及びPピクチャが、ビデオバッファ2からビデオデコーダ5に転送されて復号され、再配列バッファ6に蓄積される。

【0019】その後PTS（プレゼンテーションタイムスタンプ）の示す表示時刻になると、ビデオストリームのうちのBピクチャ、オーディオストリーム、サブピクチャストリームが、それぞれビデオバッファ2、オーディオバッファ3、グラフィックバッファ4からビデオデコーダ5、オーディオデコーダ7、グラフィックデコーダ8に転送されて復号及び表示される。このとき、再配列バッファ6に蓄積されたIピクチャ及びPピクチャも同時に表示される。

【0020】STC生成部9では、入力ストリーム中のSCR（システムクロックリフレンス）から、基準時刻であるSTC（システムタイムクロック）の時刻合わせが行われる。

【0021】STC_de1ta算出部10では、シームレス再生すべき2つのクリップのうちの先のクリップC1中のビデオストリームのうちの最後のフレームのPTSの値PTS1end、後のクリップC2中のビデオストリームのうちの最初のフレームのPTSの値PTS2start、クリップC1中のビデオストリームのうちの最後のフレームの表示期間Tppから、次のようにして、クリップC1とクリップC2との時間基準の差であるSTC_de1taが算出される。

$$STC_de1ta = PTS1end - PTS2start + Tpp$$

【0022】加算器11からは、STC生成部9で生成されたSTCとSTC_de1ta算出部10で算出されたSTC_de1taとを加算した出力STC'が得られる。

【0023】このDVR-STDは、1つのクリップの再生時には、そのクリップ中のSCRにより時刻合わせを行なったSTCを基準時刻として、P-STDと全く同様に動作する。

【0024】しかし、シームレス再生時には、クリップC1のビデオデータとクリップC2のビデオデータとを

連続的に表示するために、クリップC 1からクリップC 2への移行期間に、図のスイッチSW1, SW2, SW3等により基準時刻をSTCとSTC'との間で切り替える。例えばこのDVR-STDへのパケット多重化ストリームの入力タイミングは、次のようにして決定される。

(a) クリップC 1が入力され終えてクリップC 2の入力が開始される時刻までは、クリップC 1中のSCRにより時刻合わせを行なった基準時刻STCによって決定される。

【0025】(b) クリップC 2の入力が開始される時刻から、クリップC 1中の最後のビデオデータが表示され終える時刻まで(即ちクリップC 1からクリップC 2への移行期間)は、基準時刻がSTC'に切り替えられ、クリップC 2中のSCRによりSTCの時刻合わせが行われる。

【0026】(c) クリップC 1中の最後のビデオデータが表示され終える時刻になると、STCがその時刻でのSTC'にリセットされ、基準時刻がSTCに切り替えられる。

【0027】次に、本発明の特徴となるオーディオバッファ3の容量について、このDVR-STDと前述のE-STDとを対比しつつ説明する。E-STDの場合には、図9にも示したようにクリップの境界とストリームの境界とが一致しないことから、先のクリップC 1及び後のクリップC 2から分離されたビデオストリーム、オーディオストリームがそれぞれビデオバッファ2、オーディオバッファ3に送られる過程でのビデオストリームとオーディオストリームとの時間軸上での位置関係を図4Aと関連付けて示すと、図5Aの通りになる。

【0028】同図において、時刻t aは、クリップC 1とクリップC 2との境界であると共に、2つのビデオストリームの境界であり、時刻t bは、2つのオーディオストリームの境界である。ビデオストリームのうちの時刻t aのビデオデータとオーディオストリームのうちの時刻t bのオーディオデータとは、同期再生すべき関係にある。オーディオストリームのうちの時刻t aからt bまでの期間Tに送られる部分が、オーディオスキューである。尚、R_{video}, R_{audio}はそれぞれビデオデータからのビデオストリーム、オーディオストリームの読み出しレートである。

【0029】図4Bは、この過程でのビデオバッファの状態を例示するものである。本例では、ビデオバッファへのデータの入力レートR_{IN}とビデオバッファからのデータの出力レートR_{OUT}とが共にR_{video}と一致しているので、ビデオバッファのデータ蓄積量は一定になっていている。また、時刻t aのビデオデータは、時刻t bのオーディオデータと同期再生させるために、期間Tの間ビデオバッファに蓄積されている。(尚、MPEG2システム規格では、同一データがP-STDのバッファに1

秒を超えて蓄積されることを禁じているので、期間Tの長さは1秒が上限となる。)

【0030】図4Cは、この過程でのオーディオバッファの状態を例示するものである。本例では、オーディオバッファへのデータの入力レートR_{IN}とオーディオバッファからのデータの出力レートR_{OUT}とが共にR_{audio}と一致しているので、オーディオバッファのデータ蓄積量は一定になっている。

【0031】これに対し、図3のDVR-STDの場合には、図1にも示したようにクリップの境界とストリームの境界とが一致することから、先のクリップC 1及び後のクリップC 2からデマルチプレクサ1により分離されたビデオストリーム、オーディオストリームがそれぞれビデオバッファ2、オーディオバッファ3に送られる過程でのビデオストリームとオーディオストリームとの時間軸上での位置関係を図4Aと関連付けて示すと、図5Aの通りになる。

【0032】この場合には、時刻t aまでのビデオストリームとオーディオストリームとの関係は、E-STDの場合におけると同じになる。しかし時刻t aからは、クリップC 1中の残りのオーディオストリーム(図4Aの期間Tのオーディオスキューに相当する部分であり、図1ではクリップの右端から突出した部分)がR_{video}+R_{audio}のレートで図3のオーディオバッファ3に送られ、ビデオバッファ2にはビデオストリームが送られない。そして、そのオーディオストリームの送付が完了した時刻t cが、クリップC 1とクリップC 2との境界になる。

【0033】時刻t cから時刻t bまでは、クリップC 2中のビデオストリーム(図1でクリップの左端から突出した部分)がR_{video}+R_{audio}のレートで図3のビデオバッファ2に送られ、オーディオバッファ3にはオーディオストリームが送られない。そして、時刻t bからのビデオストリームとオーディオストリームとの関係は、E-STDの場合におけると同じになる。

【0034】期間Tのうち、時刻t aから時刻t cまでの期間T_Aと、時刻t cから時刻t bまでの期間T_Vとの内訳は、次の通りになる。

$$T_A = T \times R_{audio} \div (R_{video} + R_{audio})$$

$$T_V = T \times R_{video} \div (R_{video} + R_{audio})$$

【0035】図5Bは、この過程でのビデオバッファ2の状態を、図4Bと関連付けて示すものである。時刻t aから時刻t cまでの期間T_Aにはデータの入力がないのでデータ蓄積量が減少するが、時刻t aからt bまでの期間Tに必要となるデータは時刻t aまでに入力済みなので、この期間Tの間にビデオバッファ2がアンダーフローを起こすことはない。そして時刻t cから時刻t bまでの期間T_Vに、出力レートR_{OUT}=R_{video}よりも大きな入力レートR_{IN}=R_{video}+R_{audio}でデータが入力されることにより、時刻t bには、E-STDの

場合におけると同じデータ蓄積量にまで増加している。【0036】図5Cは、この過程でのオーディオバッファ3の状態を、図4Cと関連付けて示すものである。時刻t_aから時刻t_cまでの期間T_Aには、出力レートR_{OUT}=R_{Audio}。よりも大きな入力レートR_{IN}=R_{Video}+R_{Audio}でデータが入力されるので、データ蓄積量がE-STDの場合におけるよりも増加する。そして時刻t_cから時刻t_bまでの期間T_Vにはデータの入力がないので、時刻t_bには、E-STDの場合におけると同じデータ蓄積量にまで減少している。

【0037】従って、ビデオバッファ2の容量はE-STDの場合におけると同じであってよいのに対し、オーディオバッファ3の容量は、オーバーフローを起こさないようにするために、E-STDの場合におけるよりも増大させなければならない。その増大すべき量の最小値は、データの最大蓄積量である図5Cの時刻t_cでのデータ蓄積量と時刻t_aでのデータ蓄積量との差E_{Audio}である。この差E_{Audio}の大きさは、次のようにして算出される。

【0038】

$$\begin{aligned} E_{\text{Audio}} &= (R_{\text{IN}} - R_{\text{OUT}}) \times T_A \\ &= \{ (R_{\text{Video}} + R_{\text{Audio}}) - R_{\text{Audio}} \} \\ &\quad \times T \times R_{\text{Audio}} \div (R_{\text{Video}} + R_{\text{Audio}}) \\ &= T \times R_{\text{Video}} \times R_{\text{Audio}} \div (R_{\text{Video}} + R_{\text{Audio}}) \end{aligned}$$

【0039】そこで、図3のDVR-STDでは、オーディオバッファ3として、このE_{Audio}だけ容量を増大させたもの（図4Cに示したオーディオバッファのデータ蓄積量をO_{Audio}とするとO_{Audio}+E_{Audio}の容量のもの）を用いている。これにより、このDVR-STDでは、オーディオバッファ3をオーバーフローされることなくシームレス再生が行われる。

【0040】図6は、このE_{Audio}の大きさとR_{Audio}の大きさとの関係を、R_{Video}+R_{Audio}=10Mbps, 15Mbpsのときについてそれぞれグラフ化して示したものである。

【0041】また図7は、パケット多重化ストリーム中のオーディオストリームが、MPEGオーディオまたはドルビーAC-3により圧縮されたもの（圧縮オーディオ）である場合、PCMにより符号化されたものである場合のそれについて、R_{Video}+R_{Audio}=5, 10, 15MbpsのときのE_{Audio}の好適な大きさの具体例を示したものである。

【0042】この図7のうち、圧縮オーディオの場合のR_{Audio}の大きさ384kbps, 448kbpsは、それぞれMPEGオーディオ、ドルビーAC-3の最大ビットレートである。またPCMの場合のR_{Audio}の大きさ768kbpsは、量子化ビット数12, チャンネル数2のときサンプリング周波数32kHzに相当し（32×12×2=768）、1536kbpsは、量

子化ビット数16, チャンネル数2のときサンプリング周波数48kHzに相当する（48×16×2=1536）。

【0043】尚、以上の例では、オーディオバッファ3の容量を、オーバーフローを起こさないようにするための最小値E_{Audio}だけ増大させている。しかし、オーディオバッファ3の容量をこの値E_{Audio}よりも大きい量だけ増大させてもよいことはもちろんである。

【0044】また、以上の例では、オーディオバッファ3の容量を増大させている。しかし、時刻t_aから時刻t_bまでのオーディオバッファ3のデータ蓄積量の増加量とビデオバッファ2のデータ蓄積量の減少量とは等しいので、E-STDの場合におけるビデオバッファとオーディオバッファとの容量の総和に等しい容量の単一のバッファを設け、そのバッファのうちのビデオバッファ2として割り当てる領域とオーディオバッファ3として割り当てる領域とを時刻t_aから時刻t_bにかけて変化させるソフトウェア処理を行なうようにしてもよい。

【0045】また、以上の例では、デジタルビデオテープレコーダ用のシステムターゲットデコーダに本発明を適用しているが、それ以外のシステムターゲットデコーダ（例えばDVD用のシステムターゲットデコーダ）に本発明を適用してもよい。また、本発明は、以上の例に限らず、本発明の要旨を逸脱することなく、その他様々な構成をとりうることはもちろんである。

【0046】

【発明の効果】以上のように、本発明によれば、ストリームの接続部分で発生するオーディオスキューをそれに応じてビデオストリームと離れないよう多重化したストリームのシームレス再生を行なうシステムターゲットデコーダにおいて増大すべきオーディオバッファの容量を規定したことにより、このシステムターゲットデコーダに、オーディオバッファをオーバーフローさせることなくシームレス再生を行わせることができる。

【図面の簡単な説明】

【図1】クリーンブレークの説明に供給する図である。

【図2】クリーンブレークにおけるクリップの境界とストリームの境界とを示す図である。

【図3】本発明を適用したシステムターゲットデコーダの構成の一例を示すブロック図である。

【図4】E-STDの場合のビデオストリームとオーディオストリームとの時間軸上での関係等を示すタイミングチャートである。

【図5】図3のシステムターゲットデコーダの場合のビデオストリームとオーディオストリームとの時間軸上の関係等を示すタイミングチャートである。

【図6】図3のオーディオバッファの容量の増大量と読み出しレートとの関係を示す図である。

【図7】図3のオーディオバッファの容量の増大量の具体例を示す図である。

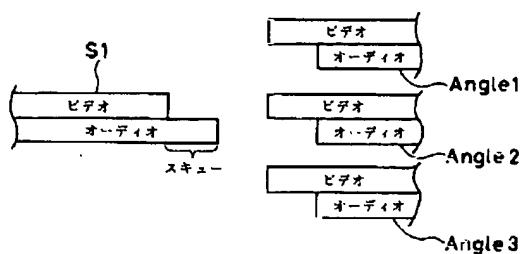
【図8】従来のパケット多重記録方式の説明に供給する図である。

【図9】従来のパケット多重記録方式におけるクリップの境界とストリームの境界とを示す図である。

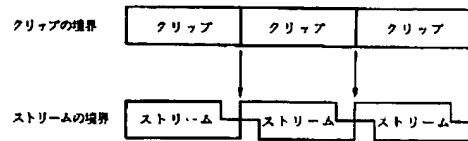
【符号の説明】

1…デマルチプレクサ、 2…ビデオバッファ、 3…オーディオバッファ、 4…グラフィックバッファ、 5…ビデオデコーダ、 6…再配列バッファ、 7…オーディオデコーダ、 8…グラフィックデコーダ、 9…STC生成部、 10…STC_delta算出部

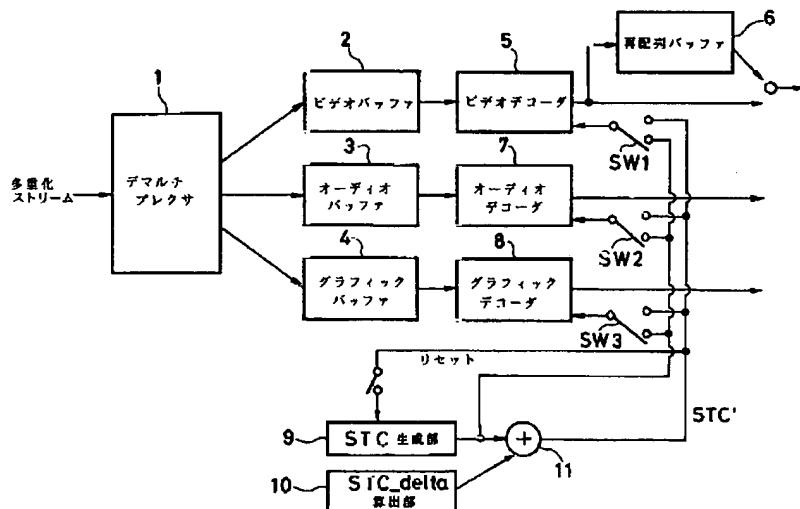
【図1】



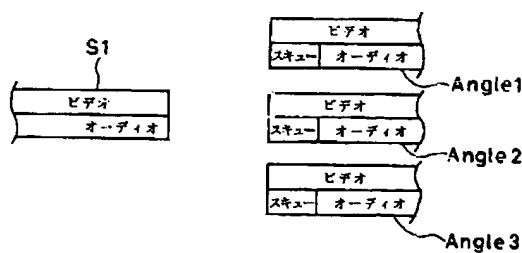
【図2】



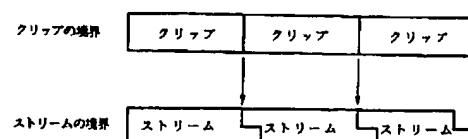
【図3】



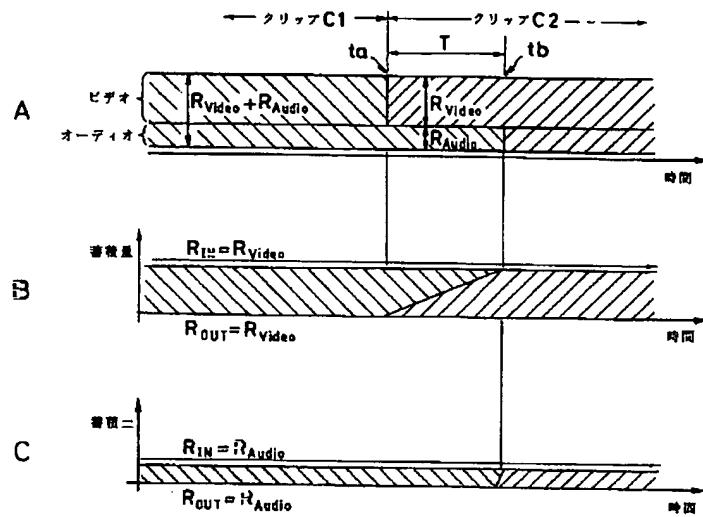
【図8】



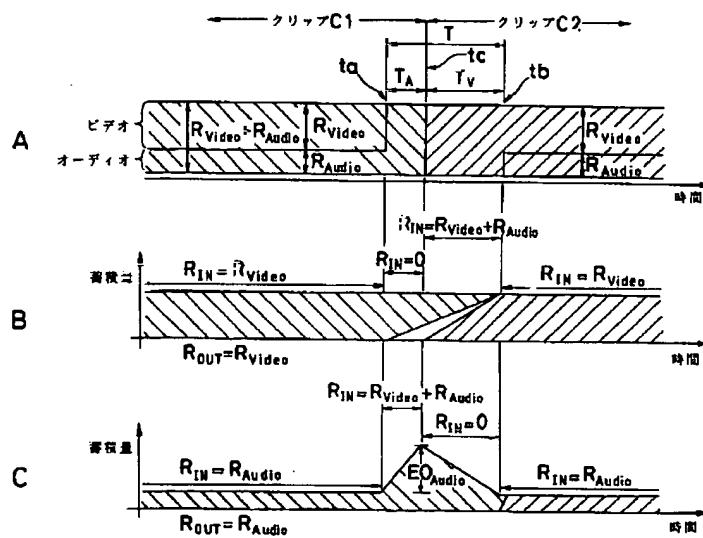
【図9】



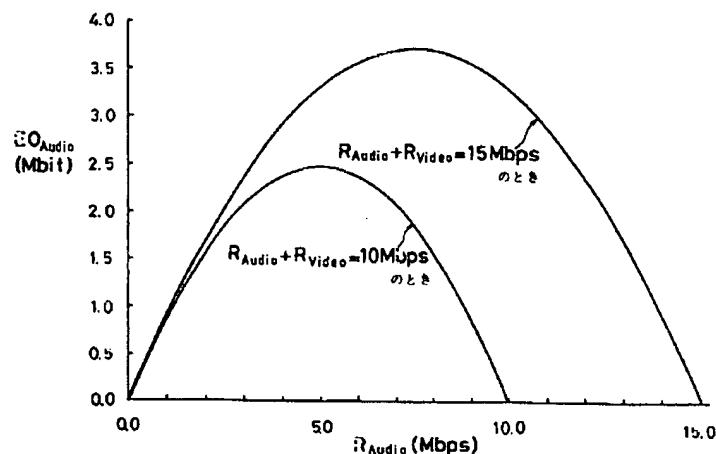
【図4】



【図5】



【図6】



【図7】

圧縮オーディオの場合

$R_{Video} + R_{Audio}$ (Mbps)	R_{Audio} (kbps)	EO_{Audio} (Mbit)
5	384	0.355
	448	0.408
10	384	0.369
	448	0.428
15	384	0.374
	448	0.435

PCMの場合

$R_{Video} + R_{Audio}$ (Mbps)	R_{Audio} (kbps)	EO_{Audio} (Mbit)
5	768	0.650
	1536	1.064
10	768	0.709
	1536	1.300
15	768	0.729
	1536	1.379